

Extended Kalman Filter-based Learning of Interval Type-2 Intuitionistic Fuzzy Logic System

Imo Eyoh, *Student Member, IEEE*^{1,2}, Robert John, *Senior Member, IEEE*^{1,2} and Geert De Maere²

¹Laboratory for Uncertainty in Data and Decision Making (LUCID)

²Automated Scheduling, Optimisation and Planning (ASAP)

School of Computer Science

University of Nottingham, UK

Email: {ije, rij, gdm}@cs.nott.ac.uk

Abstract—Fuzzy logic systems have been extensively applied for solving many real world application problems because they are found to be universal approximators and many methods, particularly, gradient descent (GD) methods have been widely adopted for the optimization of fuzzy membership functions. Despite its popularity, GD still suffers some drawbacks in terms of its slow learning and convergence. In this study, the use of decoupled extended Kalman filter (DEKF) to optimize the parameters of an interval type-2 intuitionistic fuzzy logic system of Tagagi-Sugeno-Kang (IT2IFLS-TSK) fuzzy inference is proposed and results compared with IT2IFLS gradient descent learning. The resulting systems are evaluated on a real world dataset from Australia's electricity market. The IT2IFLS-DEKF is also compared with its type-1 variant and interval type-2 fuzzy logic system (IT2FLS). Analysis of results reveal performance superiority of IT2IFLS trained with DEKF (IT2IFLS-DEKF) over IT2IFLS trained with gradient descent (IT2IFLS-GD). The proposed IT2IFLS-DEKF also outperforms its type-1 variant and IT2FLS on the same learning platform.

I. INTRODUCTION

Recent studies have shown that the application of artificial intelligence (AI) based models in prediction and forecasting provides superior performance in terms of prediction accuracy and generalisation ability compared to traditional modeling techniques [1]. Among the AI approaches, the fusion of fuzzy logic (FL) and artificial neural network (ANN) have been widely adopted especially in the area of prediction and forecasting, system identification, control and pattern recognition [2]–[5]. The two approaches - FL and ANN are known to be universal approximators [6], [7] that can identify and approximate any nonlinear system to any arbitrary degree of accuracy. The integration of FL and ANN merges the advantages of both approaches in a synergistic manner in terms of the generalisation and learning capability of ANN and the ability of the FL system to simultaneously and effectively handle uncertainties and imprecise information; and to approximately reason with these information.

The concept of a fuzzy set (FS) - a type-1 fuzzy set - was introduced by Zadeh [8] as a generalisation of the classical sets. Due to the precise nature of the membership function of FS and its inability to handle many forms of uncertainties, Zadeh [9] introduced a higher order fuzzy set, known as the type-2 fuzzy set (T2FS), as a generalisation of ordinary

type-1 FS with membership functions that are themselves fuzzy. The introduction of T2FS turns out to be a promising alternative for handling uncertainties in many applications than its type-1 counterpart [10], whose membership functions are completely certain. In classical fuzzy logic it is often the case that the degree of non-belonging of an element to a set is complementary to the degree of belonging. Thus, the membership and non-membership function of an element to a set is modeled by a single index (the membership function) value such that their sum is equal to 1. Atanassov [11] in 1986 introduced the concept of an intuitionistic fuzzy set, where the membership and non-membership value of an element to a set are evaluated separately, with some additional degree of hesitancy such that the sum of membership and non-membership function values of an element to a set lie between 0 and 1 inclusive.

Many methods have been used to optimise the parameters of fuzzy logic systems such as gradient descent (GD), simulated annealing (SA), genetic algorithm (GA), least squares method (LS) and more. Among these methods, GD, (a widely used method for neural network training) [12] are often used to tune the parameters of fuzzy systems [2], [13]–[17]. Despite the extensive use of the GD method for fuzzy systems' parameter tuning, it still suffers some drawbacks. Gradient descent being a first order derivative-based method has the disadvantage of slow convergence and may be ineffective at finding good solutions [18]. In this study we investigate the learning of interval type-2 intuitionistic fuzzy logic system of Takagi-Sugeno-Kang (IT2IFLS-TSK) fuzzy inferencing using the extended Kalman filter (EKF). However, using the standard EKF is computationally intensive because of the high dimensionality of the parameters. In order to reduce the computational burden, a simplified version of the EKF called the decoupled extended Kalman filter (DEKF) proposed in [18], [19] is used as suggested in [16] and being a second order derivative-based method, convergence is expected to be faster [16], [20]. The assumption for DEKF is that the intra-correlation among parameters of the model is high while the inter-correlation is low [20]. Hence, by decoupling the parameters and ignoring these inter-correlation [20], parameter interactions are made to occur only at the second-order level [21]. Thus, instead of

having one large vector of parameters, smaller groups (vectors) of parameters are utilised with small interactions between groups, thereby increasing the computational efficiency of the DEKF.

The DEKF has been used to train recurrent neural networks [18] and radial basis neural networks [16] with application to three different control applications and iris classification problem respectively and has been shown to provide good performance. In [20], the DEKF is used to train multilayer perceptron (MLP) networks for forecasting zonal locational marginal price (LMP) of electric energy of the Pennsylvania-New Jersey-Maryland (PJM) electricity market. Khanesar *et al* [21] used the decoupled EKF to train IT2FLS and applied the system to identification and control problems. To the best knowledge of the authors, this is the first study where a decoupled EKF is used to update the parameters of interval type-2 intuitionistic fuzzy logic system (IT2IFLS).

The organisation of the paper is as follows: in section II, we define intuitionistic fuzzy sets (IFS), type-2 intuitionistic fuzzy sets (T2IFS) and interval type-2 intuitionistic fuzzy sets (IT2IFS). Section III covers the designed system - interval type-2 intuitionistic fuzzy logic system (IT2IFLS) and the parameter update procedure is in Section IV. Section V, discusses experimental analysis and simulation results and conclusion is drawn in section VI.

II. DEFINITIONS

A. Intuitionistic Fuzzy Set (IFS)

Definition 1. [11] An IFS A^* in X can be expressed as: $A^* = \{(x, \mu_{A^*}(x), \nu_{A^*}(x)) : x \in X\}$, where $\mu_{A^*}(x) : X \rightarrow [0, 1]$ and $\nu_{A^*}(x) : X \rightarrow [0, 1]$ are the degrees of membership and non-membership of element $x \in X$ respectively such that $0 \leq \mu_{A^*}(x) + \nu_{A^*}(x) \leq 1$.

Obviously, when $\nu_{A^*}(x) = 1 - \mu_{A^*}(x)$ for every $x \in X$, then the set A is a fuzzy set.

The hesitation (also called intuitionistic fuzzy index) associated with IFS is expressed as: $\pi_{A^*}(x) = 1 - (\mu_{A^*}(x) + \nu_{A^*}(x))$.

B. Type-2 IFS (T2IFS)

Definition 2. A type-2 IFS (T2IFS) [2], \tilde{A}^* is comprised of a type-2 membership function $\mu_{\tilde{A}^*}(x, u)$, and a type-2 non-membership function $\nu_{\tilde{A}^*}(x, u)$ and expressed as:

$$\tilde{A}^* = \{(x, u), \mu_{\tilde{A}^*}(x, u), \nu_{\tilde{A}^*}(x, u) \mid \forall x \in X, \forall u \in J_x^\mu, \forall u \in J_x^\nu\}$$

where $0 \leq \mu_{\tilde{A}^*}(x, u) \leq 1$ and $0 \leq \nu_{\tilde{A}^*}(x, u) \leq 1$, such that $\forall u \in J_x^\mu$ and $\forall u \in J_x^\nu$,

$$J_x^\mu = \{(x, u) : u \in [\underline{\mu}_{\tilde{A}^*}(x), \bar{\mu}_{\tilde{A}^*}(x)]\}$$

and

$$J_x^\nu = \{(x, u) : u \in [\underline{\nu}_{\tilde{A}^*}(x), \bar{\nu}_{\tilde{A}^*}(x)]\}$$

When the membership and non-membership functions are all precise, a T2IFS collapses to IFS.

Alternatively,

$$\int_{x \in X} \left[\int_{u \in J_x^\mu} \int_{u \in J_x^\nu} \{\mu_{\tilde{A}^*}(x, u), \nu_{\tilde{A}^*}(x, u)\} \right] / (x, u)$$

When $\mu_{\tilde{A}^*}(x, u) = 1$ and $\nu_{\tilde{A}^*}(x, u) = 1$, an interval type-2 intuitionistic fuzzy set (IT2IFS) is obtained, where $\mu_{\tilde{A}^*}(x, u)$ is secondary membership and $\nu_{\tilde{A}^*}(x, u)$ is secondary non-membership functions respectively.

C. Interval Type-2 IFS (IT2IFS)

Definition 3. [22] An IT2IFS, \tilde{A}^* , can be represented by lower and upper membership functions ($\underline{\mu}_{\tilde{A}^*}(x)$, $\bar{\mu}_{\tilde{A}^*}(x)$) and lower and upper non-membership functions ($\underline{\nu}_{\tilde{A}^*}(x)$, $\bar{\nu}_{\tilde{A}^*}(x)$) for all $x \in X$ such that: $0 \leq \bar{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x) \leq 1$ and $0 \leq \underline{\mu}_{\tilde{A}^*}(x) + \bar{\nu}_{\tilde{A}^*}(x) \leq 1$.

The intuitionistic fuzzy indices used in this study are expressed as [2]:

$$\pi_c(x) = \max(0, (1 - (\mu_{\tilde{A}^*}(x) + \nu_{\tilde{A}^*}(x))))$$

$$\bar{\pi}_{var}(x) = \max(0, (1 - (\bar{\mu}_{\tilde{A}^*}(x) + \underline{\nu}_{\tilde{A}^*}(x))))$$

$$\underline{\pi}_{var}(x) = \max(0, (1 - (\underline{\mu}_{\tilde{A}^*}(x) + \bar{\nu}_{\tilde{A}^*}(x))))$$

where: $0 \leq \pi_c(x) \leq 1$ and $0 \leq \pi_{var}(x) \leq 1$. An IT2IFS \tilde{A}^* can be expressed as:

$$\begin{aligned} \tilde{A}^* &= \int_{x \in X} \int_{u \in J_x^\mu} \int_{u \in J_x^\nu} 1 / (x, u) \\ &= \int_{x \in X} \left[\int_{u \in J_x^\mu} \int_{u \in J_x^\nu} 1 / (u) \right] / x \end{aligned} \quad (1)$$

where x is the primary variable, and u is the secondary variable. The membership and non-membership bounding functions, together with the hesitation indices incorporated into the FOU's of IT2IFS, completely describe the uncertainty about an IT2IFS and defined as in Eqns (2) and (3) [2]

$$FOU_\mu(\tilde{A}^*) = \bigcup_{\forall x \in X} [\underline{\mu}_{\tilde{A}^*}(x), \bar{\mu}_{\tilde{A}^*}(x)] \quad (2)$$

$$FOU_\nu(\tilde{A}^*) = \bigcup_{\forall x \in X} [\underline{\nu}_{\tilde{A}^*}(x), \bar{\nu}_{\tilde{A}^*}(x)] \quad (3)$$

In contrast to a classical interval type-2 fuzzy set (IT2FS) with a single FOU, an IT2IFS is composed of two FOU's namely: membership function FOU - $FOU_\mu(\tilde{A}^*)$ and non-membership function FOU - $FOU_\nu(\tilde{A}^*)$.

III. INTERVAL TYPE-2 INTUITIONISTIC FUZZY LOGIC SYSTEM

The architecture of an IT2IFLS is the same as a type-2 fuzzy logic system (T2FLS) and consists of a fuzzifier, a rule base, an inference engine and an output processing block.

A. Fuzzification

During fuzzification, the input vector $x \in X$ is mapped into an IT2IFS \tilde{A}^* to obtain the membership and non-membership function values for each of the inputs. The antecedent parts of IT2IFS are the intuitionistic Gaussian membership and non-membership functions with uncertain standard deviation which are defined as in Eqns (4) to (7) [2].

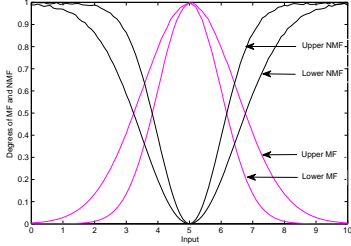


Fig. 1: IT2IFS [2]

$$\overline{\mu}_{ik}(x_i) = \exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \quad (4)$$

$$\underline{\mu}_{ik}(x_i) = \exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \quad (5)$$

$$\overline{\nu}_{ik}(x_i) = (1 - \pi_{var,ik}(x_i)) - \left[\exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{1,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \right] \quad (6)$$

$$\underline{\nu}_{ik}(x_i) = (1 - \pi_{var,ik}(x_i)) - \left[\exp\left(-\frac{(x_i - c_{ik})^2}{2\sigma_{2,ik}^2}\right) * (1 - \pi_{c,ik}(x_i)) \right] \quad (7)$$

where $\pi_{c,ik}$ and $\pi_{var,ik}$ are the IF-index of center and the IF-index of variance [23] respectively.

B. Rules

The IT2IFLS IF-THEN rule follows the same general syntax of fuzzy logic rule representation as in Eqn (8):

$$R_k : IF \ x_1 \text{ is } \tilde{A}_{1k}^* \text{ and } x_2 \text{ is } \tilde{A}_{2k}^* \text{ and } \dots \text{ and } x_n \text{ is } \tilde{A}_{nk}^* \\ THEN \ y_k = \sum_{i=1}^n w_{ik}x_i + b_k \quad (8)$$

where $\tilde{A}_{1k}^*, \tilde{A}_{2k}^*, \dots, \tilde{A}_{ik}^*, \dots, \tilde{A}_{nk}^*$ are IT2IFS and y_k is the output of the k th rule, w_{ik} 's are the consequent coefficient with offset b_k ($k = 1 \dots M$).

C. Inference

The two widely used fuzzy inferencing systems are the Mamdani and TSK-fuzzy inference. Although TSK-inferencing suffers from some form of loss of interpretability

[24], they can be more accurate than Mamdani [25] and has been used for control and regression problems with highly accurate approximations [26], [27]. Based on this premise, we adopt the modeling of IT2IFLS with TSK-fuzzy inference mechanism where the main aim is to approximate the real system as closely as possible. This paper utilises IT2IFS in the antecedent parts and crisp values in the consequent. The output of IT2IFLS is a linear combination of two Nie-Tan [28] fuzzy models (one for membership and the other for non-membership) with an additional parameter β [26] as a user defined design parameter. The final output is as defined in Eqn (9) [2].

$$y = \frac{(1 - \beta) \sum_{k=1}^M \left(\underline{f}_k^\mu + \overline{f}_k^\mu \right) y_k^\mu}{\sum_{k=1}^M \underline{f}_k^\mu + \sum_{k=1}^M \overline{f}_k^\mu} + \frac{\beta \sum_{k=1}^M \left(\underline{f}_k^\nu + \overline{f}_k^\nu \right) y_k^\nu}{\sum_{k=1}^M \underline{f}_k^\nu + \sum_{k=1}^M \overline{f}_k^\nu} \quad (9)$$

where $\underline{f}_k^\mu, \overline{f}_k^\mu$ and $\underline{f}_k^\nu, \overline{f}_k^\nu$ are the lower, upper membership and the lower, upper non-membership firing strength respectively, y_k^μ and y_k^ν are the corresponding outputs of the k th rule. The design parameter β is defined such that, $0 \leq \beta \leq 1$ and specifies the weight of membership and non-membership functions in the final output. The ‘‘prod’’ t-norm is used as the implication operator and are defined for membership function, Eqns (10) and (11) and non-membership function, Eqns (12) and (13) as follows:

$$\underline{f}_k^\mu(x) = \underline{\mu}_{\tilde{A}_{1k}^*}(x_1) \cdot \underline{\mu}_{\tilde{A}_{2k}^*}(x_2) \cdot \dots \cdot \underline{\mu}_{\tilde{A}_{nk}^*}(x_n) \quad (10)$$

$$\overline{f}_k^\mu(x) = \overline{\mu}_{\tilde{A}_{1k}^*}(x_1) \cdot \overline{\mu}_{\tilde{A}_{2k}^*}(x_2) \cdot \dots \cdot \overline{\mu}_{\tilde{A}_{nk}^*}(x_n) \quad (11)$$

$$\underline{f}_k^\nu(x) = \underline{\nu}_{\tilde{A}_{1k}^*}(x_1) \cdot \underline{\nu}_{\tilde{A}_{2k}^*}(x_2) \cdot \dots \cdot \underline{\nu}_{\tilde{A}_{nk}^*}(x_n) \quad (12)$$

$$\overline{f}_k^\nu(x) = \overline{\nu}_{\tilde{A}_{1k}^*}(x_1) \cdot \overline{\nu}_{\tilde{A}_{2k}^*}(x_2) \cdot \dots \cdot \overline{\nu}_{\tilde{A}_{nk}^*}(x_n) \quad (13)$$

where \cdot is the ‘‘prod’’ operator.

IV. PARAMETER UPDATES

A. Extended Kalman filter Parameter Update Rule

The basic idea behind the IT2IFLS prediction method is to approximate the relationship between inputs and outputs of a system as closely as possible. Assuming that the IT2IFLS model is trained by adjusting the parameters using sets of input-output pairs, then the output of a fuzzy logic system maybe represented as $y = f(X, \theta)$. The parameter X denotes the inputs into the system with θ representing the unknown parameters of the model. For IT2IFLS, these will include both the membership and non-membership functions parameters. The generic non-linear dynamic state equation can be expressed as:

$$\theta_{t+1} = f(\theta_t) + w_t \quad (14)$$

$$y_t = h(\theta_t) + v_t \quad (15)$$

where θ is the system's state, w is the process noise with zero mean and variance Q while v is the measurement noise with zero mean and variance R . For Kalman filter, the process

and measurement noise are assumed to be Gaussian and uncorrelated and:

$$\begin{aligned} E(\theta_0) &= \bar{\theta}_0, & E[(\theta_0 - \bar{\theta}_0)(\theta_0 - \bar{\theta}_0)^T] &= P_0 \\ E(w_t) &= 0, & E(w_t w_t^T) &= Q \delta_{tl} \\ E(v_t) &= 0, & E(v_t v_t^T) &= R \delta_{tl} \end{aligned}$$

where $E(\cdot)$ is the expectation operator and δ_{tl} is the Kronecker delta. The state can be estimated using Taylor expansion as:

$$\begin{aligned} f(\theta_t) &= f(\hat{\theta}_t) + F_t(\theta_t - \hat{\theta}_t) + H.O.T \\ h(\theta_t) &= h(\hat{\theta}_t) + H_t(\theta_t - \hat{\theta}_t) + H.O.T \end{aligned} \quad (16)$$

where:

$$F_t = \left. \frac{\partial f(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}_t} \quad \text{and} \quad H_t^T = \left. \frac{\partial h(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}_t}$$

and H.O.T is the higher order term. The system in Eqn (16) can be approximated as in Eqn (17) when the higher order terms are neglected.

$$\begin{aligned} \theta_{t+1} &= F_t \theta_t + w_t + \phi_t \\ y_{t+1} &= H_t^T \theta_t + v_t + \varphi_t \end{aligned} \quad (17)$$

where:

$$\begin{aligned} \phi_t &= f(\hat{x}_t) - F_t \hat{x}_t \\ \varphi_t &= h(\hat{x}_t) - H_t \hat{x}_t \end{aligned}$$

Using EKF is computationally expensive in the order of AB^2 where A is the output dimension of the dynamical system and B is the total number of parameters. Thus, for an IT2IFLS with p inputs and M number of rules and an output, the total number of parameters to be tuned is $6p + 2M(p + 1)$. The computational expense of EKF for IT2IFLS is therefore $36p^2 + 4M^2(p^2 + 2p + 1) + 24pM(p + 1)$.

1) *Decoupled Extended Kalman Filter - DEKF*: Two stages are involved in the parameter update namely: the time update and the measurement update. During the time update, the current state is projected forward in time in order to obtain a prior estimate that is used for the next step. During the measurement update, a new measurement is propagated in order to obtain the posteriori estimate. In using the DEKF to learn parameters of IT2IFLS, the antecedent and the consequent parameters are grouped into two vectors - one for the antecedent and the other for the consequent parameters. The generic parameter update rule in the i_{th} group is as in Eqns (18) to (20):

$$\theta_t^i = \theta_{t-1}^i + K_t^i [z_t - h(\theta_{t-1}^i)] \quad (18)$$

$$K_t^i = P_t^i H_t^i [(H_t^i)^T P_t^i H_t^i + R^i]^{-1} \quad (19)$$

$$P_{t+1}^i = P_t^i - K_t^i P_t^i (H_t^i)^T + Q^i \quad (20)$$

where K is the Kalman gain, P is the covariance matrix of the state estimation error, R is the measurement noise covariance and Q is the covariance of process noise. For the IT2IFLS, the unknown parameters in the antecedent are gathered into

the first vector and represented as:

$$\theta^1 = [c_{11}, c_{21}, \dots, c_{nN}, \sigma_{11}, \sigma_{21}, \dots, \sigma_{nN}]^T \quad (21)$$

where n is the number of inputs and N is the number of interval type-2 intuitionistic fuzzy partitions. The parameters of the consequent are grouped into the second vector and represented as:

$$\theta^2 = [w_{11}, w_{21}, \dots, w_{Mn}, b_1, b_2, \dots, b_M]^T \quad (22)$$

where M is the number of rules, with the membership and non-membership functions having separate Kalman parameters. The derivative matrix, H , is defined as:

$$H^1 = \frac{\partial y}{\partial \theta^1} \quad \text{and} \quad H^2 = \frac{\partial y}{\partial \theta^2} \quad (23)$$

$$(24)$$

for antecedent and consequent parameters respectively.

The update rule for the parameters in θ^1 and θ^2 then follow the same recursive procedures as in Eqns (18) to (20). By using the DEKF algorithm, the computational burden of EKF is reduced in the order $36p^2 + 4M^2(p^2 + 2p + 1)$. The computational cost of DEKF to standard EKF is in the ratio:

$$\frac{36p^2 + 4M^2(p^2 + 2p + 1)}{36p^2 + 4M^2(p^2 + 2p + 1) + 24pM(p + 1)}$$

This is a significant improvement compared to the standard EKF for training IT2IFLS.

B. Gradient Descent Learning of IT2IFLS

In order to compare our developed model, we constructed another learning procedure based on gradient descent (GD) algorithm where the GD method is used to adapt the parameters of the IT2IFLS. The cost function is defined as:

$$E = \frac{1}{2} (y^a - y)^2 \quad (25)$$

where y^a is the measured output and y is the predicted output. The GD optimisation of the IT2IFLS parameters are as follows:

$$w_{ik}(t+1) = w_{ik}(t) - \gamma \frac{\partial E}{\partial w_{ik}}$$

$$b_k(t+1) = b_k(t) - \gamma \frac{\partial E}{\partial b_k}$$

$$c_{ik}(t+1) = c_{ik}(t) - \gamma \frac{\partial E}{\partial c_{ik}}$$

$$\sigma_{1,ik}(t+1) = \sigma_{1,ik}(t) - \gamma \frac{\partial E}{\partial \sigma_{1,ik}}$$

$$\sigma_{2,ik}(t+1) = \sigma_{2,ik}(t) - \gamma \frac{\partial E}{\partial \sigma_{2,ik}}$$

where γ is the learning rate(step size), w, b are consequent parameters defined for both membership and non-membership functions and c, σ_1, σ_2 are antecedent parameters.

V. EXPERIMENTS AND RESULTS

In this section, the evaluation of the proposed learning algorithm of IT2IFLS is conducted using a real world dataset namely Australia's New South Wales (NSW) electricity price data in the year 2008. The performance of the model is analysed using the root mean square error (RMSE) and the mean absolute error (MAE) which are computed as in Eqns (26) and (27) respectively:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y^a - y)^2} \quad (26)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y^a - y| \quad (27)$$

where N is the number of test samples, y^a and y are as defined in Eqn (25). The data for the analysis have been normalised to a small range of [0,1].

A. Dataset Description

The proposed EKF-based learning IT2IFLS model is evaluated using a real world datasets from the Australia's National Electricity Market (NEM) namely New South Wales (NSW) electricity market. Similar to [29], the NSW electricity market for the year 2008 is used for the analysis. The dataset was downloaded from [30] and consists of 17568 instances with attributes of regional reference price (RRP) as the input. The price data are treated as time series data and are partitioned into four separate datasets according to [29] as representatives of the four seasons in Australia. The input data for analysis is generated from four previous values $[x(t-4), x(t-3), x(t-2), x(t-1)]$ with $x(t+1)$ as the output. There are a total of 336 data samples for each season which reduces to 331 after input generation. The first 231 data points are used for training while the remaining 100 data samples are used for testing in each season. There are 16 rules generated with a total of $6(4) + 2*16(4+1) = 184$ parameters.

B. Comparison of Results

The performance of the new learning algorithm of IT2IFLS using DEKF is evaluated on two fronts namely: (1) performance comparison with another learning algorithm such as gradient descent and (2) performance comparison with other fuzzy models trained with DEKF such as intuitionistic fuzzy logic system (IFLS) and interval type-2 fuzzy logic system - TSK (IT2FLS-TSK). The performance of each of the training algorithms was computed over 10 simulations. Figures 2 and 3 show a single season scenario between the actual and the predicted outputs of IT2IFLS-DEKF and IT2IFLS-GD with the corresponding prediction errors.

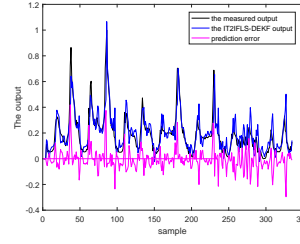


Fig. 2: Price pred. in Winter with IT2IFLS-DEKF

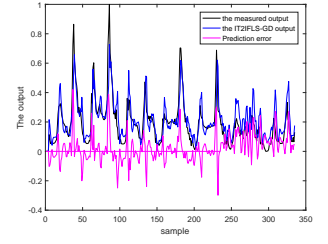


Fig. 3: Price prediction in Winter with IT2IFLS-GD

As shown in Tables I and II, IT2IFLS-DEKF exhibits superior performance over IT2IFLS-GD. We conjecture that this could be as a result of the EKF-based algorithm's ability to overcome local minima problems and to account for inter-dependence between outputs at each iterations. In Table I, the DEKF-based IFLS, IT2FLS and IT2IFLS for autumn season are very close in their modeling error. The IFLS performs slightly better than the type-2 models for the autumn season. This shows that a type-1 FLS can model uncertainty and non-linearity to some degree [21], [31]. In the overall, Tables I and II show that IT2IFLS performs better than both IFLS and IT2FLS trained with the same DEKF algorithm with reduced RMSE and MAE. Hence, using IT2IFLS can be a preferred option for handling uncertainty in many real world applications.

VI. CONCLUSION

In this study, a new learning algorithm for IT2IFLS using decoupled EKF is proposed and evaluated. The IT2IFLS possesses extra degrees of freedom which enables it to handle more uncertainties with the capacity to model non-linear input-output relationships well. From the results presented, there is a clear performance improvement of IT2IFLS trained with the decoupled EKF over the one trained with gradient descent in terms of the generalisation ability and prediction error. In general, IT2IFLS performs better than IFLS and IT2FLS with similar training apparatus on the Australia's NSW electricity market price prediction. We intend to extend this work to learning IT2IFLS with non-derivative based methods such as particle swarm optimisation (PSO) and simulated annealing (SA) and compare their performances. We also intend to explore other membership functions such as triangular and trapezoidal membership functions in the design of the proposed model.

ACKNOWLEDGEMENT

This research work is funded by the Government of Nigeria under the Tertiary Education Trust Fund (TETFund).

REFERENCES

- [1] M. Paliwal and U. A. Kumar, "Neural networks and statistical techniques: A review of applications," *Expert systems with applications*, vol. 36, no. 1, pp. 2-17, 2009.

| Period | | Summer | | | | Autumn | | | |
|--------|-----|-----------|-------------|------------|--------------|-----------|-------------|------------|--------------|
| Model | | IFLS-DEKF | IT2FLS-DEKF | IT2IFLS-GD | IT2IFLS-DEKF | IFLS-DEKF | IT2FLS-DEKF | IT2IFLS-GD | IT2IFLS-DEKF |
| RMSE | Trn | 0.0229 | 0.0243 | 0.0243 | 0.0225 | 0.0889 | 0.0891 | 0.0896 | 0.0871 |
| | Tst | 0.1112 | 0.2284 | 0.1599 | 0.0979 | 0.0407 | 0.0410 | 0.0789 | 0.0409 |
| MAE | | 0.0315 | 0.0683 | 0.0502 | 0.0284 | 0.0164 | 0.0161 | 0.0393 | 0.0167 |

TABLE I: Performance of different models and algorithms during Summer and Autumn seasons

| Period | | Winter | | | | Spring | | | |
|--------|-----|-----------|-------------|------------|--------------|-----------|-------------|------------|--------------|
| Model | | IFLS-DEKF | IT2FLS-DEKF | IT2IFLS-GD | IT2IFLS-DEKF | IFLS-DEKF | IT2FLS-DEKF | IT2IFLS-GD | IT2IFLS-DEKF |
| RMSE | Trn | 0.0836 | 0.0846 | 0.0916 | 0.0791 | 0.0715 | 0.0754 | 0.0802 | 0.0723 |
| | Tst | 0.0439 | 0.0429 | 0.0553 | 0.0422 | 0.0960 | 0.0954 | 0.1333 | 0.0821 |
| MAE | | 0.0182 | 0.0184 | 0.0239 | 0.0175 | 0.0342 | 0.0367 | 0.0477 | 0.0335 |

TABLE II: Performance of different models and algorithms during Winter and Spring seasons

- [2] I. Eyoh, R. John, and G. De Maere, "Interval type-2 intuitionistic fuzzy logic system for non-linear system prediction," in *2016 IEEE International Conference on Systems, Man and Cybernetics*, 2016, pp. 1063–1068.
- [3] R. H. Abiyev and O. Kaynak, "Type 2 fuzzy neural structure for identification and control of time-varying plants," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 12, pp. 4147–4159, 2010.
- [4] H. A. Hagras, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, pp. 524–539, 2004.
- [5] A. Khosravi, S. Nahavandi, D. Creighton, and D. Srinivasan, "Interval type-2 fuzzy logic systems for load forecasting: A comparative study," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1274–1282, 2012.
- [6] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [7] B. Kosko, "Fuzzy systems as universal approximators," *IEEE transactions on computers*, vol. 43, no. 11, pp. 1329–1333, 1994.
- [8] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.
- [9] —, "The concept of a linguistic variable and its application to approximate reasoning-i," *Information Sciences*, vol. 8, pp. 199–249, 1975.
- [10] J. M. Mendel, "Uncertain rule-based fuzzy logic system: introduction and new directions," 2001.
- [11] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.
- [12] D. Psaltis, A. Sideris, and A. A. Yamamura, "A multilayered neural network controller," *IEEE control systems magazine*, vol. 8, no. 2, pp. 17–21, 1988.
- [13] F. Guély and P. Siarry, "Gradient descent method for optimizing various fuzzy rule bases," in *Fuzzy Systems, 1993., Second IEEE International Conference on*. IEEE, 1993, pp. 1241–1246.
- [14] J. Kim and N. Kasabov, "Hyfis: adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems," *Neural Networks*, vol. 12, no. 9, pp. 1301–1319, 1999.
- [15] A. Habbi and M. Zemat, "An improved self-tuning mechanism of fuzzy control by gradient descent method," in *Proceedings of the 17th European Simulation Multiconference, Nottingham, UK*, 2003.
- [16] D. Simon, "Training fuzzy systems with the extended kalman filter," *Fuzzy sets and systems*, vol. 132, no. 2, pp. 189–199, 2002.
- [17] J. R. Castro, O. Castillo, P. Melin, and A. Rodríguez-Díaz, "A hybrid learning algorithm for a class of interval type-2 fuzzy neural networks," *Information Sciences*, vol. 179, no. 13, pp. 2175–2193, 2009.
- [18] G. V. Puskorius and L. A. Feldkamp, "Neurocontrol of nonlinear dynamical systems with kalman filter trained recurrent networks," *IEEE Transactions on neural networks*, vol. 5, no. 2, pp. 279–297, 1994.
- [19] —, "Decoupled extended kalman filter training of feedforward layered networks," in *Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on*, vol. 1. IEEE, 1991, pp. 771–777.
- [20] Y. Ma, P. B. Luh, K. Kasiviswanathan, and E. Ni, "A neural network-based method for forecasting zonal locational marginal prices," in *Power Engineering Society General Meeting, 2004. IEEE*. IEEE, 2004, pp. 296–302.
- [21] M. A. Khanesar, E. Kayacan, M. Teshnehlab, and O. Kaynak, "Extended kalman filter based learning algorithm for type-2 fuzzy logic systems and its experimental evaluation," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4443–4455, 2012.
- [22] D. D. Nguyen, L. T. Ngo, and L. T. Pham, "Interval type-2 fuzzy c-means clustering using intuitionistic fuzzy sets," in *IEEE Third World Congress on Information and Communication Technologies (WICT)*, 2013, pp. 299–304.
- [23] P. Hájek and V. Olej, "Intuitionistic fuzzy neural network: The case of credit scoring using text information," in *Engineering Applications of Neural Networks*. Springer, 2015, pp. 337–346.
- [24] M. J. Gacto, R. Alcalá, and F. Herrera, "Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures," *Information Sciences*, vol. 181, no. 20, pp. 4340–4360, 2011.
- [25] M. J. Gacto, M. Galende, R. Alcalá, and F. Herrera, "Metsk-hd e: A multiobjective evolutionary algorithm to learn accurate tsf-fuzzy systems in high-dimensional and large-scale regression problems," *Information Sciences*, vol. 276, pp. 63–79, 2014.
- [26] M. B. Began, W. W. Melek, and J. M. Mendel, "Parametric design of stable type-2 tsf fuzzy systems," in *IEEE Annual Meeting of the North American Fuzzy Information Processing Society, (NAFIPS)*, 2008, pp. 1–6.
- [27] A. H. Sonbol, M. S. Fadali, and S. Jafarzadeh, "Tsf fuzzy function approximators: Design and accuracy analysis," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 3, pp. 702–712, 2012.
- [28] M. Nie and W. W. Tan, "Towards an efficient type-reduction method for interval type-2 fuzzy logic systems," in *Fuzzy Systems, 2008. FUZZ-IEEE 2008. (IEEE World Congress on Computational Intelligence)*. IEEE International Conference on. IEEE, 2008, pp. 1425–1432.
- [29] P. Areekul, T. Senjyu, N. Urasaki, and A. Yona, "Neural-wavelet approach for short term price forecasting in deregulated power market," *Journal of International Council on Electrical Engineering*, vol. 1, no. 3, pp. 331–338, 2011.
- [30] <https://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/>, accessed: 2017-02-03.
- [31] R. Sepulveda, O. Castillo, P. Melin, O. Montiel, and L. T. Aguilar, "Evolutionary optimization of interval type-2 membership functions using the human evolutionary model," in *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*. IEEE, 2007, pp. 1–6.